Identification of Piecewise Linear Models of Complex Dynamical Systems

Ronald L. Westra, **Mihály Petreczky** **Ralf L.M. Peeters**

DKE, Maastricht University, The Netherlands,
{Westra,M.Petreczky,Ralf.Peeters}@maastrichtuniversity.nl

Abstract: The paper address the realization and identification problem for a subclass of piecewise-affine hybrid systems. The paper provides necessary and sufficient conditions for existence of a realization, a characterization of minimality, and an identification algorithm for this subclass of hybrid systems. The considered system class and the identification problem are motivated by applications in systems biology.

Keywords: Realization theory, minimization, identification, hybrid systems, network topology, gene-protein networks.

1. INTRODUCTION

In this paper we address the realization and identification problem for a subclass of piecewise-affine hybrid systems.

Contribution of the paper We define the class of *piecewise-linear systems* (abbreviated by **PWL**). **PWL**s are a subclass of piecewise-affine hybrid systems. The continuous dynamics of **PWL** is determined by a finite collection of affine subsystems. However, in contrast to traditional piecewise-affine systems, we allow any change of the continuous state during a discrete-state transition, as long as the new state belongs to the set of designated initial states of the affine subsystem associated with the new discrete state. In addition, we do not impose any specific mechanism for triggering discrete-state transitions.

We formulate the realization problem for this system class and partially solve it by providing necessary and sufficient conditions for existence of a realization. We also present conditions for minimality. We show that the outputs of any **PWL** can also be described by a switched AR model. The main conclusion is that any system can be transformed to a minimal system *with one discrete state* while preserving input-output behavior. This means that without further restrictions, the identification problem for such systems is not necessarily interesting. We discuss a number of restrictions on the system structure which avoid this problem.

Note that the conclusion above is not valid for other classes of hybrid systems. For hybrid systems from Petreczky and van Schuppen (2010), there examples of input-output maps which provenly cannot be realized by systems with one discrete state.

In addition, we present an identification algorithm for systems with full observations. This algorithm is illustrated by example of physical and biological relevance.

Motivation The motivation for studying realization theory for **PWL**s is that it provides the theoretical foundations for systems identification. The motivation for investigating identification of **PWL**s is the following.

- First, several systems of interest can be modeled by **PWL**s and **PWL**s tend to be convenient for analysis. However the parameters of **PWL** models are often not directly available and hence they have to be estimated from measurements.
- The problem of estimating the network dynamics of complex biological systems can be reduced to the identification problem for **PWL**s.

Identification of PWLs and biological networks Below we will elaborate on the relationship between identification of **PWL**s and the estimation of dynamic interactions of complex dynamical systems.

Numerous vital processes in nature involve complex signaling networks, varying from gene-protein interaction networks Westra et al. (2007) and complex communication in microbes Gera and Srivastava (2006) to the synchronization in the heart of higher animals Heijman et al. (2009). In most cases, these signaling networks do not have a fixed topology, but vary their structure according to their internal states and certain external conditions Adami et al. (2000). This flexibility allows the organism to exhibit complex behavior and flexible responses to various environmental conditions.

The underlying biological problem is to describe these flexible dynamical networks, based on (partial) observations of their internal states. Examples of potential observations are time series of gene expressions, protein densities, oxygen stress and sugar concentrations.

In almost all cases, there is no explicit mathematical model available that even remotely describes the network dynamics. This is caused by the sheer size of the problem; the many thousands of genes, RNAs, proteins and other types of molecules involved, the complexity and idiosyncrasy of their interactions, and the large amount of noise in biological systems. Basically, many agents involved in the network are unknown let alone their mechanism of interaction.

Mathematically, the processes above can be viewed as nonlinear dynamical systems with unknown parameters. Such systems often have several equilibria, and hence their behavior can be seen as a mixture of certain affine systems, where each affine system is obtained by linearizing the original system around one of the equilibria. That is, such systems can be approximated by a **PWL**. In addition to their simplicity, the advantage of **PWL** approximations is that they neatly capture the interaction among various state components around various equilibria. More precisely, interaction between the ith and jth state components can be viewed as the property that the (i, j)th entry of the Jacobian of the system around equilibrium is nonzero. Since we are often interested exactly in interactions rather than the detailed dynamics, it makes sense to recast the problem of identifying such interactions into the problem of identifying **PWL**s approximations. While the paper does not directly address identification of interactions, we believe that the presented results represent a step towards the solution of that problem.

Related work Identification of piecewise-linear (-affine) hybrid systems has been subject of intensive research, Vidal et al. (2002); Ma and Vidal (2005); Bako et al. (2009b); Paoletti et al. (2007); Vidal (2008); Ferrari-Trecate et al. (2003); Juloski et al. (2004); Bako et al. (2009a); Roll et al. (2004); Fox (2009). Realization theory for hybrid systems was investigated in Weiland et al. (2006); Paoletti et al. (2010); Petreczky (2006); Petreczky and van Schuppen (2010); Grossman and Larson (1995). The application of hybrid systems to modelling complex dynamical systems, originating from biology, and their identification has been a subject of intensive research Casey et al. (2005); de Jong (2002); Porreca et al. (2010); Koutroumpas et al. (2007); Cinquemani et al. (2008).

The results of this paper are new, to the best of our knowledge. What distinguishes the contribution of this paper from the existing work is **(a)** the system class considered, in particular, the emphasis on continuous-time systems, **(b)** the emphasis on realization theory and algorithms, **(c)** the details of the identification algorithm.

Outline of the paper §2 presents the formal definition of the system class of interest. §3 discusses the relationship between nonlinear systems with complex dynamics and piecewise-affine hybrid systems. §4 presents the results on realization theory, and §5 presents the identification algorithm and the results of the numerical experiments.

Notation We use the standard notation. We denote by $T = [0, +\infty)$ the time-axis. We denote by I_n the $n \times n$ identity matrix. We denote the set of natural numbers including zero by $\mathbb{N}$.

2. PIECEWISE-LINEAR SYSTEMS (PWL)

The aim of the section is to define the class of piecewise-linear systems formally.

Definition 1. (PWL). A *piecewise-linear system* (abbreviated as PWL) is a dynamical system determined by

$$\Sigma \begin{cases} \dot{x}(t) = A_{q(t)}x(t) + a_{q(t)} \\ y(t) = C_{q(t)}x(t) + c_{q(t)} \\ x(t^+) \in \mathcal{X}_{q(t^+),0} \end{cases} \tag{1}$$

Here $Q = \{1, \ldots, D\}$ is the finite set of discrete modes, $\mathbb{R}^p$ is the output space, $\mathbb{R}^{n_q}$ is the state-space of the system in mode $q \in Q$, For each $q, \in Q$ $A_q \in \mathbb{R}^{n_q \times n_q}$, $a_q \in \mathbb{R}^{n_q}$, $C_q \in \mathbb{R}^{p \times n_q}$, and $c_q \in \mathbb{R}^p$ are the parameters of the affine system in mode $q \in Q$, $\mathcal{X}_{q,0} \subseteq \mathbb{R}^n$ – is the set of initial states of the affine system in mode $q \in Q$. The *state space* $\mathcal{H}_\Sigma$ of Σ is $\mathcal{H}_\Sigma = \bigcup_{q \in Q}\{q\} \times \mathbb{R}^{n_q}$ We call Σ *linear*, if $c_q = 0$ and $a_q = 0$ for all $q \in Q$, otherwise Σ is called *affine*.

We will use the following short-hand notation

$$\Sigma = (p, Q, \{(n_q, A_q, a_q, C_q, c_q, \mathcal{X}_{q,0}) \mid q \in Q\}).$$

Informally, the evolution of Σ takes place as follows. As long as the value of the discrete state $q(t)$ at time t does not change, the continuous state and the continuous output an time t change according to the affine system $\dot{x}(t) = A_{q(t)}x(t) + a_{q(t)}$ and $y(t) = C_{q(t)}x(t) + c_{q(t)}$. The discrete state can change at any time, however, we do not allow consecutive changes of discrete states immediately one after the other. If the discrete state changes to $q(t^+)$ at time t, then the new continuous state should satisfy $x(t^+) \in \mathcal{X}_{q(t^+),0}$, i.e. it should belong to the set of the designated initial states of the new discrete state $q(t^+)$. Note that we do not specify the mechanism which triggers the change of discrete state. We also do not specify the initial discrete state, it is chosen by an unspecified mechanism. Hence, the *description above allows for several state- and output-trajectories.*

Note that the definition of **PWL**s presented above differs from the one used in the literature. In the standard definition it is assumed that after a discrete-state transition, the new continuous state depends on the previous one. In contrast, here we only require that the new continuous state belongs to the designated set of initial states of the affine system associated with the new discrete state. This implies that *the behavior of the system in discrete state q depends only on the affine system associated with that discrete state and it does not depend on the discrete modes which were visited in the past.* Another consequence of the definition is that the model is not predictive. That is, the knowledge of the system parameters and the inputs does not uniquely determine the state- and output-trajectory of the system.

In order to define the evolution of **PWL**s formally, we need to introduce the following notation and terminology.

Definition 2. (Collins (2005)). A time event sequence is a strictly monotone sequence $(t_n)_{n=0}^{n^*}$ such that $n^* \in \mathbb{N} \cup \{+\infty\}$, $t_0 = 0$ and for all $0 < n < n^*$, $0 \leq t_n < t_{n+1}$. If $n^* = +\infty$ then we require that $\sup\{t_n \mid n \in \mathbb{N}\} = +\infty$. If $n^* < +\infty$, then by convention $t_{n^*+1} = +\infty$.

The role of time event sequences is to formalize the time instances at which discrete events occur. The restrictions formulated in the definition imply that the set of switching times does not have an accumulation point, i.e. no Zeno-behavior can take place.

Definition 3. (State-trajectory). A state-trajectory of Σ is a map $\xi : T \to \mathcal{H}_\Sigma$ such that there exists a time event sequence $(t_i)_{i=0}^{n^*}$ and a sequence of discrete modes $(q_i \in Q)_{i=0}^{n^*}$ such that for all $0 \leq i \leq n^*$, $i \in \mathbb{N}$, it holds that for all $s \in [t_i, t_{i+1})$, $\xi(s) = (q_i, x(s - t_i))$ and

$$\dot{x}(t) = A_{q_i}x(t) + a_{q_i} \text{ and } x(0) \in \mathcal{X}_{q_i,0}.$$

The time event sequence $(t_n)_{n=0}^{n^*}$ is called the sequence of *switching times* of the state-trajectory ξ.

We *denote by* $BS(\Sigma)$ *the set of all state-trajectories of* Σ.

Definition 4. (Output-trajectory). An output-trajectory of Σ is a map $y : T \to \mathbb{R}^p$ such that the following holds. There exists a state-trajectory ξ of Σ, such that $y(t) = \upsilon_\Sigma(x(t))$ for all $t \in T$. Here, υ_Σ is the *readout-map* of Σ, defined as

$$\upsilon_\Sigma : \mathcal{H}_\Sigma \ni (q, x) \mapsto C_q x \in \mathbb{R}^p.$$

We *denote by* $B(\Sigma)$ *the set of all output-trajectories of* Σ.

In the sequel, unless stated otherwise, f *denotes a function* $f : T \to \mathbb{R}^p$. The definition above implies that the external behavior of a **PWL** is exactly a function of this type.

Definition 5. (Realization). The function f is said to be realized by **PWL** Σ, if if is an output-trajectory of Σ, i.e. if $f \in B(\Sigma)$. In this case Σ is called a realization of f.

Definition 6. The dimension of a **PWL** Σ is a tuple $(|Q|, n)$, where $n = \sum_{q \in Q} n_q$ if Σ is linear, and $n = \sum_{q \in Q}(n_q + 1)$, if Σ is affine.

That is, the first element of $\dim \Sigma$ is the number of discrete states, the second element is the number of continuous state components. The additional dimension in the case of affine **PWL**s stems from the need to store the vectors a_q, c_q, $q \in Q$. We use the following partial order relation on $\mathbb{N} \times \mathbb{N}$. We say that $(p, q) \in \mathbb{N}$ is smaller than or equal $(r, s) \in \mathbb{N}$, denoted by $(p, q) \le (r, s)$, if $p \le r$ and $q \le s$. Note that the order relation $\le$ in $\mathbb{N} \times \mathbb{N}$ is indeed a partial order, it is not possible to compare all elements of $\mathbb{N} \times \mathbb{N}$.

Definition 7. (Minimality). A **PWL** Σ is said to be a minimal realization of f, if for any **PWL** $\hat{\Sigma}$ which is a realization of f, $\dim \Sigma \le \dim \hat{\Sigma}$.

The definition above implies that if Σ is minimal, then it is possible to compare its dimension to the dimension of any other realization of f. Since we work with a partial order on dimensions,the existence of a minimal system is not at all obvious.

Problem 1. (Realization problem for **PWL**s). Find conditions for existence of a **PWL** realization of f. Characterize minimal **PWL**s realizations of f. Find algorithms for computing a **PWL**s realization from finite data.

As it was indicated before, our motivation for studying the realization problem for **PWL**s is to lay the theoretical foundations for identification of **PWL**s. We present below the formulation of the identification problem.

Problem 2. (Identification problem). Assume that the value of f and its derivatives up to order r, $r \ge 0$ is measured at time instances $t_1 < \ldots < t_k$. Based on the (possibly noisy) data $\{f^{(l)}(t_i)\}_{i=1\ldots,k,l=1,\ldots,r}$, find a **PWL** realization of f.

3. APPROXIMATION BY **PWLS**

Our main motivation for studying identification of **PWL**s is the following. The problem of identifying interaction networks of complex dynamical systems can be reduced to the identification problem of **PWL**s. Indeed, consider a partially observed non-linear system of the form

$$\dot{x}(t) = f(x(t)) \text{ and } y(t) = h(x(t)) \qquad (2)$$

where $f : \mathbb{R}^N \to \mathbb{R}^N$ and $h : \mathbb{R}^N \to \mathbb{R}^p$ are sufficiently smooth functions.

We are interested in the dynamics of the network structure of (2). By *the network structure of* (2) *at time* t we mean the directed graph with nodes numbered by $1, \ldots, N$ and where an edge goes from node i to node j if the ith component of $x(t)$ influences the jth component of $\dot{x}(t)$. As the values of $x(t)$ changes, so does $f(x(t))$. Hence the network structure described above depends on time.

The basic question of interest is how to derive the dynamics of the interaction network based on the observed output y. One obvious approach is to estimate f and h from y, but this is feasible only if special restrictions are put on f and h. Notice, however, that if we look at local linearizations of f around an equilibrium and the interactions are strong enough, then $x_i(t)$ influences $\dot{x}_j(t)$, if the (i, j)th entry of the Jacobian of f is non-zero.

Prompted by the considerations above we propose the following approach to modelling and identification of the network structure of systems of the form (2). We assume that (2) has finitely many equilibrium points. We assume that under influence of noise the system jumps between neighborhoods of these equilibrium points. We assume that the transition from the neighborhood of one equilibrium point to the neighborhood of another equilibrium point takes place very fast. Hence, the behavior of the system during the transition can be ignored. We associate a **PWL** Σ of the form (1) with (2) as follows. Let D be the number of equilibrium points. For each $q \in Q$, let e_q be the corresponding equilibrium point and define

$$A_q = Df(e_q) \text{ and } a_q = -A_q e_q$$
$$C_q = Dh(e_q) \text{ and } c_q = h(e_q)$$

Here Df and Dh denote the Jacobian of f and h respectively. Let $\mathcal{X}_{q,0}$ be a suitably chosen neighborhood of e_q. It then follows that if the state $x(t)$ of (2) is close enough to e_q, we can approximate (2) as follows

$$\dot{x}(t) = f(x(t)) \approx A_q x(t) + a_q$$
$$y(t) = h(x(t)) \approx C_q x(t) + c_q$$

As it was noted above, the relevant information regarding the interactions of state variables is already included in A_q. Hence, *for the purposes of network reconstructions, identification of the **PWL**s approximation is sufficient.*

4. REALIZATION THEORY OF **PWLS**

Below we present some basic results on realization theory of **PWL**s. First we present a characterization of minimality, after that we present conditions for existence of a realization and a realization algorithm. Throughout the section, f *denotes a function* $f : T \to \mathbb{R}^p$.

4.1 Minimality

In order to present our first result, we introduce the notion of a *linear system with state-jumps.*

Definition 8. A linear system with state jumps, abbreviated as (**LSSJ**) is a linear **PWL** Σ of the form (1) with one discrete state, i.e. $D = 1$. We will identify the **LSSJ** Σ with the collection of data $(n, C, A, \mathcal{X}_0)$, where $n = n_1, C = C_1, A = A_1$ and $\mathcal{X}_{1,0} = \mathcal{X}_0$.

The reason we call a **PWL** with one discrete state a linear system with state jumps is that the system behaves as a linear system, with the exception that its state occasionally jumps back to one of the initial states. For **LSSJ**s, we can easily define the concepts of observability and span-reachability.

Definition 9. A **LSSJ** $\Sigma = (n, C, A, \mathcal{X}_0)$ is called *observable*, if (C, A) is an observable pair, and Σ is called *span-reachable*, if $\mathrm{Span}\{A^k x_0 \mid k = 0, \dots, n-1, x_0 \in \mathcal{X}_0\} = n$.

Remark 1. Note that $Z = \mathrm{Span}\{A^k x_0 \mid k = 0, \dots, n-1, x_0 \in \mathcal{X}_0\}$ is an A invariant subspace containing $\mathcal{X}_0$, hence by restricting A and C to Z we can transform Σ to a span-reachable **LSSJ** Σ_r while preserving output trajectories, i.e. $B(\Sigma_r) = B(\Sigma)$. By applying linear observability reduction, we can transform the **LSSJ** Σ_r to a **LSSJ** Σ_m which is span-reachable, observable and has the same output trajectories as Σ, i.e. $B(\Sigma_m) = B(\Sigma)$.

Theorem 1. (Minimality). Assume that f admits a realization by a **PWL**. Then f admits a minimal **PWL** realization Σ such that Σ is a span-reachable and observable **LSSJ**. A **LSSJ** realization of f is minimal if and only if it is span-reachable and observable.

The theorem above says that in general, the external behavior of any **PWL**s can be represented by a linear system with several initial states. That is, without further restrictions, the realization and identification problems for **PWL**s are equivalent to that of **LSSJs**.

The proof of Theorem 1 relies on the following transformations.

Definition 10. (**PWL** to linear **PWL**). Define the *linear* **PWLs** $L(\Sigma)$ associated with an affine Σ as follows.

$$L(\Sigma) = (p, Q, \{(n_q^L, A_q^L, a_q^L, C_q^L, c_q^L, \mathcal{X}_{q,0}^L) \mid q \in Q\})$$

where $n_q^L = n_q + 1$, $c_q^L = 0$, $a_q^L = 0$ and

$$A_q^L = \begin{bmatrix} A_q & a_q \\ 0 & 0 \end{bmatrix}, \; C_q^L = [C_q \;\; c_q]$$

$$\mathcal{X}_{q,0}^L = \{(x^T, 1)^T \in \mathbb{R}^{n_q+1} \mid x \in \mathcal{X}_{q,0}\}.$$

Proposition 1. The output trajectories of Σ and $L(\Sigma)$ coincide, i.e. $B(\Sigma) = B(L(\Sigma))$, and $\dim \Sigma = \dim L(\Sigma)$.

Proof. [Sketch] A state trajectory ξ of $L(\Sigma)$ is always a map of the form $\xi(t) = (q(t), (x(t)^T, 1)^T)$, $t \in T$ such that $\hat{\xi}(t) = (q(t), x(t))$, $t \in T$ is a state-trajectory of Σ. Since $\upsilon_{L(\Sigma)}((q, (x^T, 1)^T)) = C_q x + c_q = \upsilon_\Sigma((q, x))$, the statement of the proposition follows.

Definition 11. (Linear **PWL** to **LSSJ**). Let Σ be a linear **PWL** of the form (1). Define the **LSSJ** $LS(\Sigma)$ associated with Σ as follows.

$$LS(\Sigma) = (n, C, A, \mathcal{X}_0),$$

where $n = \sum_{q \in Q} n_q$ and

$$A = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \cdots & A_D \end{bmatrix} \text{ and } C = [C_1 \; \cdots \; C_D]$$

$$\mathcal{X}_0 = \bigcup_{q \in Q} \hat{\mathcal{X}}_{q,0}$$

$$\hat{\mathcal{X}}_{q,0} = \{(\underbrace{0, \dots, 0}_{q-1-times}, x, 0, \dots, 0)^T \mid x \in \mathcal{X}_{q,0}\}, \forall q \in Q$$

Proposition 2. The output trajectories of Σ and $LS(\Sigma)$ are the same, i.e. $B(\Sigma) = B(LS(\Sigma))$, moreover, $\dim LS(\Sigma) = (1, n)$ where $(|Q|, n) = \dim \Sigma$.

Proof. [Sketch] For each $q \in Q$, let $\mathcal{X}_q$ be the subset of $\mathbb{R}^n$ of the form $z = (\underbrace{0, 0, \dots, 0}_{q-1-times}, x^T, 0, \dots, 0)^T$, $x \in \mathbb{R}^{n_q}$. For each $z \in \mathcal{X}_q$ of the above form, let $\Pi_q(z) = x \in \mathbb{R}^{n_q}$. It then follows that $\xi(t) = (1, x(t))$ is a state-trajectory of $LS(\Sigma)$ if and only if $\hat{\xi}(t) = (q(t), \Pi_{q(t)}(x(t)))$ is a state-trajectory of Σ, where $q(t) = q$ for some $q \in Q$ if and only if $x(t) \in \mathcal{X}_q$. Since $\upsilon_{LS(\Sigma)}((1, x)) = C_q \Pi_q(x) = \upsilon_\Sigma((q, \Pi_q(x))$ for all $x \in \mathcal{X}_q$, the statement of the proposition follows.

Using the two transformations defined above, and Remark 1, we can present the proof of Theorem 1.

Proof. [Proof of Theorem 1] Assume that Σ is a **PWL** realization of f. If Σ is affine, then replace Σ with the associated linear **PWL** $L(\Sigma)$, which is also a realization of f. Hence, we can assume that Σ is already a linear **PWL**. Construct then the **LSSJ** $LS(\Sigma)$ associated with Σ and apply Remark 1 to $LS(\Sigma)$ to obtain a span-reachable and observable **LSSJ** Σ_m such that $B(\Sigma_m) = B(LS(\Sigma)) = B(\Sigma)$. It then follows that $\dim \Sigma_m \leq \dim LS(\Sigma) \leq \dim \Sigma$.

Hence, since Σ was arbitrary, it is enough to look for minimal realizations among span-reachable and observable **LSSJ**s. Notice that for any two **LSSJ**s Σ_1 and Σ_2, either $\dim \Sigma_1 \leq \dim \Sigma_2$ or $\dim \Sigma_2 \leq \dim \Sigma_1$, hence, among all the possible **LSSJ**s realizations of f, there must exist a minimal one.

4.2 Existence of a realization

The conditions for existence of a realization will be formulated using the rank of the Hankel-matrix of f. In order to define the Hankel-matrix of f, we have to define the notion of Markov-parameters. To that end, we need the notion of *piecewise-analytic functions.*

Definition 12. (Piecewise-analytic). The map f is called *piecewise-analytic*, if there exist a finite or infinite number of time instances, $t_i \in T$, $t_i < t_{i+1}$, $i \leq N_f$, $i \in \mathbb{N}$ for some $N_f \in \mathbb{N} \cup \{+\infty\}$, such that the following holds. If $N_f < +\infty$, let $t_{N_f+1} = +\infty$. Then we require that $t_0 = 0$ and $\bigcup_{i=0}^{N_f} [t_i, t_{i+1}) = T$ and for each $i \in \mathbb{N}$, $i \leq N_f$, f is analytic on $[t_i, t_{i+1})$, but f is not analytic on any neighborhood of t_i in $\mathbb{R}$. We call the points $\{t_i\}_{i=0}^{N_f}$ the points of *non-analyticity.* We define the set

$$I_f = \{i \in \mathbb{N} \mid i \leq N_f\}$$

of indices of points of non-analyticity.

The intuition behind the definition is as follows. If f has a realization by a **PWL** Σ, then the only points where f is not analytic are the points where the corresponding state-trajectory of Σ switches from one discrete mode to another. In fact, one can show that there always exists a state-trajectory of Σ which yields f as output trajectory and which switches only at time instances at which f is not analytic. Hence, if f has a realization by a **PWL**, then f is piecewise-analytic and the points of non-analyticity tell us the switching times of a **PWL** realization of f.

In the sequel, *f is assumed to be piecewise-analytic.*

Definition 13. (Markov-parameters). Assume that f is piecewise-analytic and let $\{t_i\}_{i=0}^{N_f}$ be the points of non-analyticity of f. For each $i \in I_f$, define the ith *Markov-parameter* $\mathbf{M}_i^f$ of f as a sequence $\mathbf{M}_i^f : \mathbb{N} \to \mathbb{R}^p$

$$\forall k \in \mathbb{N} : \mathbf{M}_i^f(k) = \frac{d^k}{dt^k} f(t_i + s)|_{s=0}.$$

It is easy to see that the collection of Markov-parameters $\{\mathbf{M}_i^f\}_{i=0}^{N_f}$ determines the map f uniquely. We use the Markov-parameters to define the Hankel-matrix of f.

Definition 14. (Hankel-matrix). We define the *Hankel-matrix* H_f of f as the infinite matrix, rows of which are indexed by $\mathbb{N} \times \{1, \dots, p\}$, and columns of which are indexed by $\mathbb{N} \times I_f$. The entry of H_f indexed by row index (i, r) and by column index (j, l) equals

$$[H_f]_{(i,r),(j,l)} = (\mathbf{M}_l^f(i + j))_r$$

where $(\mathbf{M}_l^f(i + j))_r$ denotes the rth entry of Markov-parameter $\mathbf{M}_l^f(i+j)$. The rank of H_f, denoted by rank H_f, is the dimension of the linear space spanned by the columns of H_f.

Theorem 2. (Existence of a **PWL** realization). The map f can be realized by a **PWL** if and only if it is piecewise-analytic and rank $H_f < +\infty$. Moreover, a minimal **LSSJ** realization of f of dimension $(1, \text{rank } H_f)$ can be constructed from H_f .

Proof. [Sketch] **only if** Assume that f can be realized by a **PWL**. Then it can be realized by a **LSSJ** $\Sigma = (n, C, A, \mathcal{X}_0)$. Then $M_i^f(k) = CA^k x_i \ \forall i \in I_f, k \in \mathbb{N}$ for some $x_i \in \mathcal{X}_0$, i.e. $\{M_i^f(k)\}_{k=0}^{\infty}$ are the Markov-parameters of the linear system (C, A) from some initial condition. It then follows from linear systems theory that rank $H_f \leq n < +\infty$.

if Assume that n = rank H_f and fix a basis in the linear span of the columns of H_f. Define C as the matrix in this basis of the linear map which maps a column to the vector formed by the rows of that column indexed by $(0, 1), \dots, (0, p)$, in this order. Define the matrix A as the matrix in this basis of the linear map which maps the column indexed by (j, l) to the column indexed by $(j + 1, l)$. Finally, let $\mathcal{X}_0$ be the set of coordinate vectors of the columns of H_f indexed by $(0, l)$, $l \in I_f$. Then $\Sigma_f = (n, C, A, \mathcal{X}_0)$ is a **LSSJ** realization of f. If Σ is another **LSSJ** realization of f, then from the **only if** part it follows that rank $H_f = \dim \Sigma_f \leq \dim \Sigma$, i.e. Σ_f is a minimal realization of f.

In Algorithm 1 we present a Kalman-Ho-like realization algorithm for **LSSJ**s and hence **PWL**s. To this end, we need the following definition.

Definition 15. (Finite Hankel sub-matrix). Fix integers $R, L, M > 0$. Define the set

$$I_f^R = \{i \in I_f \mid i \leq R\}.$$

Define the submatrix $H_{f,L,M,R}$ of H_f as the matrix which is formed by the intersection of the rows of H_f indexed by the elements of $I_L = \{0, \dots, L\} \times \{1, \dots, p\}$ and the columns of H_f indexed by the elements of $J_{M,R} = \{0, \dots, M\} \times I_f^R$.

Algorithm 1
Inputs: Hankel-matrix $H_{f,L,M+1,R}$.
Output: LSSJ $\Sigma_{L,M,R}$.

```
Compute the decomposition $H_{f,L,M+1,R} = \mathbf{OR}$ such that $\mathbf{O} \in \mathbb{R}^{I_L \times n}$ and $\mathbf{R} \in \mathbb{R}^{n \times J_{M+1,R}}$ and rank $\mathbf{R}$ = rank $\mathbf{O} = n$.
Define $\bar{\mathbf{R}} \in \mathbb{R}^{n \times J_M}$ as the matrix formed by the columns of $\mathbf{R}$ indexed by elements of $J_{M,R}$. Let $\hat{\mathbf{R}} \in \mathbb{R}^{n \times J_{M,R}}$ the matrix such that the column of $\hat{\mathbf{R}}$ indexed by $(j, l)$ equals the column of $\mathbf{R}$ indexed by $(j + 1, l)$.
Define $\Sigma_{L,M,R} = (n, C, A, \mathcal{X}_0)$ as follows.
    • The $i$th row of $C$ equals the row of $\mathbf{O}$ indexed by $(i, 0)$, $i = 1, \dots, p$.
    • The matrix $A$ is the solution of the equation $\hat{\mathbf{R}} = A\bar{\mathbf{R}}$
    • The set $\mathcal{X}_0$ consists of the columns of $\mathbf{R}$ indexed by indices of the form $(0, l)$, $l \in I_f^R$.
```

Theorem 3. (Correctness of Algorithm 1). If rank $H_f \leq n$, and $|\{M_i^f \mid i \in I_f\}| < +\infty$, then for some $R > 0$, the **LSSJ** returned by Algorithm 1 is a minimal realization of f. The conditions above hold if f has a realization by a **LSSJ** of dimension at most n and with a finite set of initial states.

4.3 Ill-posedness of the realization problem

The results on realization theory of **PWL**s indicate that the realization and identification problems for **PWL**s are in general ill-posed in the following sense. Realizability by a **PWL**s is equivalent to realizability by a minimal **PWL**s with one single discrete state. Hence, there seems to be no intrinsic way to choose discrete states based on the data. Since the original intention was use the combination of the dynamics in various discrete modes to explain complex dynamics, the above conclusion is an unpleasant one.

Note that the main problem is that one can increase the dimension of the continuous state-space to encode discrete states. One way to remedy this is to place restriction on the dimension. To this end, we introduce the notion of $K - N$ realization.

Definition 16. ($K - N$). A **PWL** Σ is said to be a $K - N$ **PWL** if Σ is of the form (1) and $|Q| \leq K$ and for all $q \in Q$, $n_q \leq N$. The map f is said to admit a $K - N$ realization, if there exists a $K - N$ **PWL** which is a realization of f.

The motivation for using $K - N$ realizations is that by choosing an appropriate N, we can make sure that the growth in the number of continuous states cannot be used to replace discrete states. Below we present conditions for existence of a $K - N$ realization of f. To this end, we have to introduce the following concept.

Definition 17. (Hankel-matrices & partition). Consider the partitioning $\mathbf{C} = (\mathbf{C}_q)_{q=1}^K$ of the set of I_f. For each $q = 1, \dots, K$, let $H_{f,\mathbf{C}}^q$ be the sub-matrix of Hankel-matrix H_f which is formed by the columns of H_f indexed by indices of the form (j, l), $j \in \mathbb{N}$ and $l \in \mathbf{C}_q$.

Theorem 4. (Existence of $K - N$ realizations). The map f has a realization by a linear $K - N$ **PWL** if and only if there exists a partitioning $\mathbf{C} = (\mathbf{C}_i)_{i=1}^D$ of I with $D \leq K$

such that for all $i = 1, \ldots, D$, rank $H^i_{f,\mathbf{C}} \le N$. If the latter condition holds, then Σ can be computed from H_f with $Q = \{1, \ldots, D\}$ and with $n_q =$ rank $H^q_{f,\mathbf{C}}$ for all $q \in Q$.

Theorem 4 above says that f can be realized by a $K - N$ realization if the columns of the Hankel-matrix H_f can be divided in K clusters such that for each fixed $l \in I$, all columns indexed by (i, l) end up in the same cluster. Moreover, the linear span of the elements of each cluster should have dimension at most N.

In order to prove Theorem 4, we need the following transformation of a **LSSJ** to a $K - N$ realization.

Definition 18. (**LSSJ** to **PWL**). Consider a partition $\mathbf{C} = (\mathbf{C}_l)_{l=1}^K$ of I_f for some $K > 0$. Consider an observable **LSSJ** $\Sigma_l = (n, C, A, \mathcal{X}_0)$ such that Σ_l is a realization of f. Then there exists a unique state trajectory ξ of Σ_l, such that the switching times of ξ coincide with the points of non-analyticity $(t_i)_{i=0}^{N_f}$ of f and $\upsilon_{\Sigma_l}(\xi(t)) = f(t)$ for all $t \in T$. Define the **PWL** realization $PW(\Sigma_l, \mathbf{C})$ of f *associated with* Σ_l *and* $\mathbf{C}$ as follows. The **PWL** $PW(\Sigma_l, \mathbf{C})$ is of the form (1), such that $Q = \{1, \ldots, K\}$, and

- For all $q \in Q$, let $\hat{\mathcal{X}}_{q,0}$ be the set of states $x(t_l)$ such that for some $l \in \mathbf{C}_q$, $\xi(t_l) = (1, x(t_l))$. Define
$$\mathcal{X}_q = \mathrm{Span}\{A^k x \mid k = 0, \ldots, n-1, x \in \hat{\mathcal{X}}_{q,0}\}$$
Set $n_q = \dim \mathcal{X}_q$ and choose a basis of $\mathcal{X}_q$.
- For each $q \in Q$, define A_q as the matrix in the basis of $\mathcal{X}_q$ of the linear map $\mathcal{X}_q \ni x \mapsto Ax \in \mathcal{X}_q$ obtained by restricting A to $\mathcal{X}_q$.
- Let $\mathcal{X}_{q,0}$ be the set of vector representations in the basis of $\mathcal{X}_q$ of the elements of $\hat{\mathcal{X}}_{q,0}$.
- Let C_q be the matrix representation in the basis of $\mathcal{X}_q$ of the linear map $\mathcal{X}_q \ni x \mapsto Cx \in \mathbb{R}^p$ obtained by restricting C to $\mathcal{X}_q$.

Proposition 3. If Σ_l is a **LSSJ** realization of f, then the **PWL** $PW(\Sigma_l, \mathbf{C})$ is a realization of f.

Proof. [Sketch of the proof of Theorem 4] **only if** Assume that Σ is a **PWL** realization of f of the form (1), such that $|Q| \le K$ and $n_q \le N$ for all $q \in Q$. Then there exists a state trajectory ξ of Σ such that $f = \upsilon_\Sigma \circ \xi$ and the switching times of ξ contain the points of non-analyticity $(t_i)_{i=0}^{N_f}$ of f. Define $\mathbf{C}_q = \{i \in I_f \mid \xi(t_i) = (q, x_i)$ for some $x_i \in \mathcal{X}_{q,0}\}$. It then follows that $\mathbf{C} = (\mathbf{C}_q)_{q=1}^D$ is a partitioning of I_f. Moreover, for any $i \in \mathbf{C}_q$, M^f_i is the Markov-parameter of the linear system (C_q, A_q) from some initial condition. Hence, by using linear systems theory we can show that rank $H^q_{f,\mathbf{C}} \le n_q \le N$.

if If the conditions of the theorem hold, then rank $H_f = n < +\infty$. Construct the **LSSJ** realization Σ_f of f, as in the proof of Theorem 2. Apply Definition 18 to obtain a linear **PWL** $\Sigma = PW(\Sigma_f, \mathbf{C})$. Then Σ has $D \le K$ discrete states. Notice that the space $\mathcal{X}_q$ from Definition 4 is then isomorphic to the column space of $H^q_{f,\mathbf{C}}$, $q = 1, \ldots, D$. Hence $n_q = \dim \mathcal{X}_q =$ rank $H^q_{f,\mathbf{C}} \le N$.

Similarly to the general case, in Algorithm 2 we state an algorithm for computing a $K - N$ realization of f.

Theorem 5. (Correctness of Algorithm 2). If f has a $K - N$ realization of dimension at most $(K, \min\{L, M\})$, and and with a finite set of initial states, then for some $R > 0$, Algorithm 1 returns a minimal realization of f.

```
Algorithm 2
Inputs: Hankel-matrix $H_{f,L,M+1,R}$.
Output: PWL $\Sigma$
Apply Algorithm 1 to $H_{f,L,M+1,R}$ and denote the result by $\Sigma_{L,M,R}$.
Compute a partitioning $\mathbf{C} = (\mathbf{C}_i)_{i=1}^D$ of $I^R_f$ such that $D \le K$ and for all $i = 1, \ldots, D$, rank $H^i_{f,\mathbf{C},L,M,K} \le N$. Here $H^i_{f,\mathbf{C},L,M,R}$ is the sub-matrix of $H_{f,L,M,R}$ formed by the columns $(j, l)$ where $j \le M$ and $l \in \mathbf{C}_q$.
Return the PWL $\Sigma = PW(\Sigma_{L,M,R}, \mathbf{C})$.
```

4.4 *Equivalence of* **PWL** *and switched AR models*

Below we show that f is realizable by a **PWL** if and only if f satisfies a switched AR model.

Definition 19. (**SARS** models). A *switched AR system* (abbreviated as **SARS**) is a tuple
$$\mathcal{I} = (p, n, Q, \{A_{q,i} \mid q \in Q, i = 1, \ldots, n\}) \tag{3}$$
where $Q = \{1, \ldots, D\}$, $D > 0$, $n_q > 0$ and $A_{q,i} \in \mathbb{R}^{p \times p}$ for all $i = 1, \ldots, n_q$. The function f is said to satisfy the **SARS** model $\mathcal{I}$, if the following holds. Let $(t_i)_{i=0}^{N_f}$ be the points of non-analiticity of f. For any $i \in I_f$, and for any $t \in [t_i, t_{i+1})$, denote by $f^{(k)}(t)$ the kth order right-hand derivative $\frac{d^k}{ds^k} f(t+s)|_{s=0}$. Then we require that for any $i \in I_f$ there exist $q(i) \in Q$, such that $\forall t \in [t_i, t_{i+1})$,
$$f^{(n)}(t) = \sum_{k=1}^{n-1} A_{q(i),k} f^{(n-k)}(t). \tag{4}$$
The set Q is called the set of discrete modes of $\mathcal{I}$.

Notice that if $D = 1$, then the **SARS** $\mathcal{I}$ corresponds to an AR system, and f satisfies $\mathcal{I}$ if it satisfies an AR equation. Note, however, that in contrast to trajectories of an AR systems, f is not smooth.

Theorem 6. The function f has a realization by a **PWL** with D discrete states if and only if there exists a **SARS** $\mathcal{I}$ with D discrete modes such that f satisfies $\mathcal{I}$.

Proof. [Sketch] **only if** Let Σ be a realization of f. Without loss of generality we can assume that Σ is linear, it is of the form (1) and $n_q = n$ for all $q \in Q$. Then for each $i \in I_f$, there exists $q(i) \in Q$ such that on $[t_i, t_{i+1})$, f is an output trajectory of the linear system $(C_{q(i)}, A_{q(i)})$. From classical theory we then know that there exists matrices $A_{q(i),1}, \ldots, A_{q(i),n}$ such that (4) holds.

if Assume f satisfies a **SARS** $\mathcal{I}$ of the form (4). Define the linear **PWL** $\Sigma_{\mathcal{I}}$ of the form (1) such that $\Sigma_{\mathcal{I}}$ is a realization of f as follows. The sets of discrete modes of $\Sigma_{\mathcal{I}}$ and $\mathcal{I}$ coincide, $n_q = n$, $\mathcal{X}_{q,0} = \mathbb{R}^{n_q}$ and the matrices C_q and A_q, $q \in Q$ are the matrices of the linear system which corresponds to the AR $y^{(n)}(t) = \sum_{j=1}^n A_{q,j} y^{(n-j)}(t)$. Here we consider the linear system, state vector of which is the regressor $x(t) = ((y^{(n-1)})^T(t), \ldots, y^T(t))^T$.

Theorem 1 and Theorem 6 yield the following.

Corollary 1. The function f has a realization by a **PWL** if and only if it satisfies an AR system, i.e. for some $A_1, \ldots, A_n \in \mathbb{R}^{p \times p}$, $n > 0$,

$$\forall t \in T : f^{(n)}(t) = \sum_{i=1}^{n} A_i f^{(n-i)}(t).$$

5. IDENTIFICATION ALGORITHM FOR **PWLS**

Below we present an algorithm which computes a realization with full observations based on measurements at finitely many time instances. More precisely, the algorithm solves the following problem.

Problem 3. Fix integers $D > 0$ and $n > 0$. Consider a piecewise-analytic function $f : T \to \mathbb{R}^n$ and a finite sequence of time instances $t_1 < \ldots < t_M$ and assume that $\{f(t_i), \dot{f}(t_i)\}_{i=1}^{M}$ are known, i.e. we know that value of f and its derivative at time instances $t_1, \ldots, t_k$. Find a **PWL** Σ of the form (1) with full-observations such that Σ realizes f, and $C_q = I_n$, $c_q = 0$, and $n_q = n$ for all $q \in Q$.

Motivation of the identification problem The motivation for considering the identification problem with full observations is the following.

(1) Notice that the identification problem with full observations and the identification problem for **SARS**s are equivalent. Indeed, a **PWL** with full observation can be considered as a **SARS**s with $n = 2$. Conversely, for a **SARS** $\mathcal{I}$, the associated **PWL** $\Sigma_{\mathcal{I}}$ from the proof of Theorem 6 can be viewed as a **PWL** with full observations, if the high-order derivatives of f can be measured. Hence, by Theorem 6, *the identification problem for **PWL**s is equivalent to the identification problem for **PWL** with full observations*, if we assume that high-order derivatives of the output can be measured too.

(2) The problem of identification with full observations is still a non-trivial problem. Even in this case we have problems with identifiability, see Example 1.

(3) The solution of Problem 3 enables us to prove experimentally the feasibility of approximating complex systems by **PWL**s. This is done by applying the identification algorithm which solves Problem 3 to simulated trajectories of several well-known complex systems.

Example 1. Consider the **LSSJ**s $\Sigma_1 = (2, C_1, A_1, \mathcal{X}_0^1)$ and $\Sigma_2 = (2, C_2, A_2, \mathcal{X}_0^2)$ where $C_1 = C_2 = I_2$ is the identity matrix and $\mathcal{X}_0^1 = \mathcal{X}_0^2 = \{(1,0)^T\}$ and the remaining parameters are as follows; $A_1 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ and $A_2 = \begin{bmatrix} 0 & 2 \\ 0 & 3 \end{bmatrix}$. It then follows that $f(t) = (1,0)^T$, $t \in T$ can be realized both by Σ_1 and Σ_2, but clearly $\Sigma_1 \neq \Sigma_2$. In other words, *realizations of f with full observation are not identifiable.* The reason for this phenomenon is that the state of both Σ_1 and Σ_2 live in the subspace (x, y), $y = 0$, hence the difference in A_1 and A_2 is not visible.

Remark 2. (Related work on **SARS**s). There is a wealth of results on identification of **SARS**, mostly addressing the discrete-time SISO case. Many of these results, in particular, the algebraic approach Ma and Vidal (2005); Vidal (2008); Bako et al. (2009b) could probably be adapted to the system class of this paper. Investigating such adaptations remains a topic of future research. Note that it is not at all clear that the convergence results for the discrete-time case have meaningful counterparts in the continuous-time case.

In Algorithm 3 we present a solution for solving Problem 3. Algorithm 3 is an iteration consisting of the following steps. First, the algorithm starts with a random initialization of $\{A_q, a_q\}_{q \in Q}$. Subsequently, in each iteration step, the current estimates $A_q, a_q, \{w_{q,i}\}_{q \in Q, i=1,\ldots,M}$ are updated by applying first the step 3 for estimating the updated weights with the fixed system parameters $\{A_q, a_q\}_{q \in Q}$, and then applying step 4 with the previously updated weights $\{w_{q,i}\}_{q \in Q, i=1,\ldots,M}$ for updating the estimates of the system parameters $\{A_q, a_q\}_{q \in Q}$. The interpretation of the weights is as follows: $w_{q,i}$ is one if in t_i the discrete state $q \in Q$ is active, and it is zero otherwise. The algorithm terminates when an absolute criterion E falls below a pre-specified threshold ϵ. The iteration fails if after a pre-defined maximum number of iterations T_{max} the criterion E has not yet reached the lower threshold ϵ. Unfortunately, the conditions under which the algorithm terminates and returns a correct realizations are not known yet.

Remark 3. (Convergence of Algorithm 3). The criterion E of the algorithm always decreases or stays the same during the iterations, and the same holds for the cost functions in step 3–4. Hence, the criterion E and the cost functions in step 3–4 will converge to a fixed value. Whether this value is a global optimum remains a topic of future research. Note that the true parameters of a realization of f render E zero, but there might be many of them, see Example 1.

5.1 Numerical example

We conducted several numerical experiments. They were all performed on a PC with Intel Pentium 4 CPU 2.8 GHz processor and 1.21 GB RAM memory under Windows XP, using Matlab 7.5 including the optimization toolbox. In line with the definitions above, we use K and N to quantify the dimension of the **PWL** of interest and M to denote the number of data points. Dynamical data $\{f(t_i), \dot{f}(t_i)\}_{i=1}^{M}$ was sampled from a given dynamical model $\dot{x} = f(x)$ at regular time intervals $t_i = i.\Delta$, $i = 1, \ldots, M$ for some $\Delta > 0$. Several initial states were chosen randomly, and a sampled trajectory was obtained for each of them. We then concatenated these sampled trajectories into one time series. Notice that concatenation of finite components of two output trajectories of a **PWL** is itself a finite component of a valid output trajectory. Hence, the combined data can be viewed as originating from the measurements of one valid output trajectory of a **PWL**. Gaussian noise with zero mean and variance σ^2 was added to the obtained data. Finally, Algorithm 3 was applied to the resulting data. In the numerical experiments we first tested the performance of the algorithm on artificial **PWL** systems. Furthermore, we studied the algorithm on Lorenz systems and on the Tyson-Novak model for the cell cycle of budding yeast.

*Simulations on artificial **PWL** systems* We generate an artificial **PWL** system containing K affine subsystems, each of the same dimension $n_q = N$. The state switching is obtained by partitioning the state space in K subsets $\{V_1, \ldots, V_K\}$. The switching is generated by using the following switching law: the discrete mode i is active, if the continuous state belongs to V_i.

In the absence of noise, and if the subsystems are sufficiently sampled, the estimation algorithm is usually able

Algorithm 3 Identification algorithm
Inputs: data $\{f(t_i), \dot{f}(t_i)\}_{i=1}^{M}$
Output: PWL Σ

```
Initialize the estimates of $A_q, a_q$, $q \in Q$, $k := 0$.
**repeat**
    Solve the optimization for the weights
    $w = \arg\min_{w^*=\{w^*_{q,i}\}_{q\in Q, i=1,\dots,M}} E(\{A_q, a_q\}_{q\in Q}, w^*)$
    $\forall m = 1, \dots, M : \sum_{q\in Q} w^*_{q,m} = 1$, $\forall q \in Q : w^*_{q,m} \in [0,1]$
    where for $w = \{w_{q,i}\}_{q\in Q, i=1,\dots,M}$, $A_q \in \mathbb{R}^{n\times n}$, $a_q \in \mathbb{R}^n$, $q \in Q$,
        $E(\{A_q, a_q\}_{q\in Q}, w) =$
        $= \frac{1}{M}\sum_{q\in Q}\sum_{m=1}^{M} w_{q,m}||\dot{f}(t_m) - (A_q f(t_m) + a_q)||^2$
    Note that the optimization problem above is a linear programming problem, hence the optimal values of $w = \{w_{q,i}\}_{q\in Q, i=1,\dots,M}$ take values in the set $\{0,1\}$.
    Using the weights $w = \{w_{q,i}\}_{q\in Q, i=1,\dots,M}$ recalculate the optimal values of $A_q, a_q$, $q \in Q$ by solving the following minimization problem.
    $\{A_q, a_q\}_{q\in Q} = \arg\min_{A^*_q, a^*_q, q\in Q} E(\{A^*_q, a^*_q\}_{q\in Q}, w)$    (5)
    The optimization problem (5) is equivalent to the following linear least squares problem. Collect the unknowns $A_q, a_q$, $q \in Q$ into the matrix $S$
        $S = [A_1 \ a_1 \ A_2 \ a_2 \ \cdots \ A_D \ a_D]$
    Define the matrices $L \in \mathbb{R}^{(n+1)D\times M}$ and $Y \in \mathbb{R}^{n\times M}$
        $Y = \left[\dot{f}(t_1) \ \cdots \ \dot{f}(t_m)\right]$
        $L = \begin{bmatrix} w_{11}r_1 & \cdots & w_{M1}r_M \\ \vdots & \dots & \vdots \\ w_{1D}r_1 & \cdots & w_{MD}r_M \end{bmatrix}$
        $r_i = \left[f(t_i)^T \ 1\right]^T, \forall i = 1, \dots, M$
    Then the solution of (5) can be read off from the entries of $S$, where
        $S = \arg\min_{\hat{S}\in\mathbb{R}^{n\times(n+1)D}} ||Y - \hat{S}L||^2$
    Update $E = E(\{A_q, a_q\}_{q\in Q}, w)$ using the new values.
**until** $E < \epsilon$
Return $\Sigma$ of the form (1) with $n_q = n$, $C_q = I_n$, $c_q = 0$ and $A_q, a_q$ as computed in the last iteration, for all $q \in Q$, and $\mathcal{X}_{q,0} = \mathbb{R}^n$ for all $q \in Q$.
```

to find the system parameters, though occasionally it gets stuck in a local minimum. The amount of empirical data required to reconstruct the system parameters, within the bounds set by the noise, increases proportionally with N and K. Figure 1 below shows the result of the algorithm for $N = 2$, $K = 5$, $M = 1800$ data points. The Voronoi partitioning of state space of the original dataset (black) and the reconstructed dataset (red) are identical. The reconstructed system parameters and the original system parameters are identical as well.

Fig. 1.

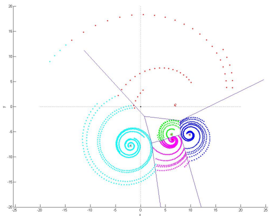


Fig. 2.

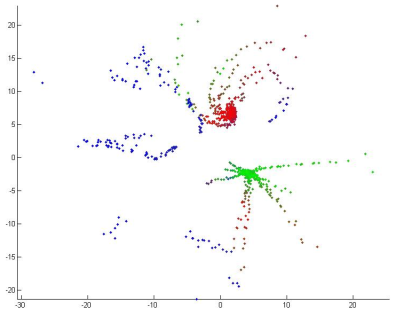
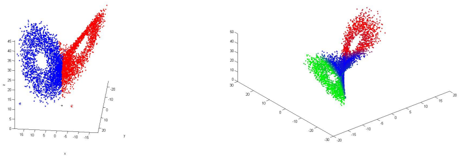


Fig. 3.



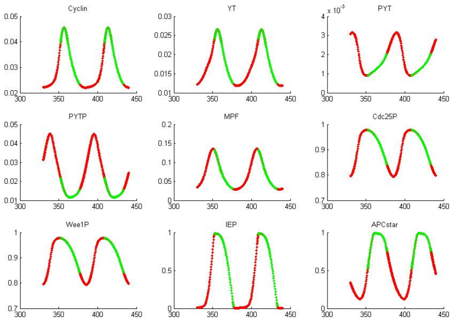
Further tests reveal that Algorithm 3 is able to find the correct parameters even if the regions $\{V_1, \dots, V_K\}$ are disconnected.

When applying zero-mean Gaussian noise $\mathcal{N}(0, \sigma^2)$, the accuracy of the reconstruction decreases as σ increases. Figure 2 below shows the result on a dataset of dimension $N = 2$ with $K = 5$ subsystems and $M = 1800$ data points with a SNR (signal-to-noise ratio) of 5%. Two subsystems could not be reconstructed, but the correlation between the three reconstructed systems with the best fitting original subsystems is 98.4

Chaotic systems We applied Algorithm 3 to Lorenz systems Lorenz (1963). This example provides insight how the identification algorithm performs on systems that are not piecewise affine by nature.

The Lorenz attractor can be visualized approximately as consisting of two linear subsystems Left and Right, with two different central attractors. A third region can be defined that involves the transition between the Left and Right subsystem. In Figure 3 the application of Algorithm 3 to Lorenz systems with $M = 5000$ data points is presented for $K = 2, 3$. For $K = 2$ the two main parts of the Lorenz attractor are found. For $K = 3$ also the transition region from the left to the right attractor is found. Applying more than 3 subsystems does not improve the performance.

Fig. 4.



Biological cell cycle models We applied Algorithm 3 to data simulated using the Tyson-Novak model Chen et al. (2004) for the yeast cell cycle. The cell cycle is the succession of events whereby a cell grows and divides into two daughter cells. The simulated data was used to construct a **PWL** approximation of the system using Algorithm 3. This resulted in a clear partition of the cell cycle dynamics as a function of the selected number of classes K. The best qualitative results were obtained for $K = 2$, see Figure 4. Moreover, for $K > 3$ the results did not improve. It was found that the obtained two discrete modes correspond to the phases S (DNA synthesis) and M (mitosis) of the cell cycle. The reconstruction remains valid under noise with SNR at most 10%.

6. CONCLUSIONS

We presented some basic results on realization theory of **PWL**s and a practical identification algorithm. Analysis of the correctness and convergence of the presented algorithm remains a topic of future research.

REFERENCES

Adami, C., Ofria, C., and Collier, T.C. (2000). Evolution of biological complexity.

Bako, L., Mercre, G., and Lecoeuche, S. (2009a). Online structured subspace identification with application to switched linear systems. *International Journal of Control*, 82, 1496–1515.

Bako, L., Mercre, G., Vidal, R., and Lecoeuche., S. (2009b). Identification of switched linear state space models without minimum dwell time. In *IFAC Symposium on System Identification, Saint Malo, France.*

Casey, R., de Jong, H., and Gouzé, J.L. (2005). Piecewise-linear models of genetic regulatory networks: Equilibria and their stability. *J. Math. Biology.*

Chen, K., Calzone, L., Csikasz-Nagy, A., Cross, F., Novak, B., and Tyson, J. (2004). Integrative analysis of cell cycle control in budding yeast. *Mol. Biol. Cell*, 15, 3841 – 3862.

Cinquemani, E., Porreca, R., Ferrari-Trecate, G., and Lygeros, J. (2008). Subtilin Production by Bacillus Subtilis: Stochastic Hybrid Models and Parameter Identification. *IEEE Trans. on Circuits and Systems-I.*

Collins, P. (2005). Hybrid trajectory spaces. Technical report, Centrum voor Wiskunde en Informatica (CWI), Amsterdam.

de Jong, H. (2002). Modeling and simulation of genetic regulatory systems: A literature overview. *Journal of Computational Biology*, 9(1), 67–103.

Ferrari-Trecate, G., Musellu, M., Liberati, D., and Morari, M. (2003). A clustering technique for the identification of piecewise-affine systems. *Automatica*, 39, 205–217.

Fox, E. (2009). *Bayesian Nonparametric Learning of Complex Dynamical Phenomena.* Ph.D. thesis, MIT, Cambridge, MA.

Gera, C. and Srivastava, S. (2006). Quorum-sensing: The phenomenon of microbial communication. *Current Science*, 90(5).

Grossman, R. and Larson, R. (1995). An algebraic approach to hybrid systems. *Theoretical Computer Science*, 138, 101–112.

Heijman, J., Westra, R., and Rudy, Y. (2009). A computational model of beta-adrenergic signaling in cardiac myocytes. *Heart Rhythm Journal*, 6(5), 97 –104.

Juloski, A., Weiland, S., and Heemels, M. (2004). A bayesian approach to identification of hybrid systems. In *Proceedings of 43rd IEEE Conference on Decision and Control.*

Koutroumpas, K., Cinquemani, E., and Lygeros, J. (2007). Randomized optimization methods in parameter identification for biochemical network models. In *FOSBE 2007.*

Lorenz, E.N. (1963)). Deterministic nonperiodic flow. *J. Atmos. Sci.*, 20, 130 – 141.

Ma, Y. and Vidal, R. (2005). Identification of deterministic switched arx systems via identification of algebraic varieties. In *Hybrid Systems: Computation and Control*, volume 3414 of *LNCS*, 449 – 465.

Paoletti, S., Juloski, A., Ferrari-Trecate, G., and Vidal, R. (2007). Identification of hybrid systems: A tutorial. *European Journal of Control*, 13(2-3), 242 – 260.

Paoletti, S., Roll, J., and Garulli, A.and Vicino, A. (2010). On the input-output representation of piecewise affine state space models. *IEEE Transactions on Automatic Control*, 55, 60 – 73.

Petreczky, M. (2006). *Realization Theory of Hybrid Systems.* Ph.D. thesis, Vrije Universiteit, Amsterdam.

Petreczky, M. and van Schuppen, J. (2010). Realization theory for linear hybrid systems. *IEEE Trans. on Automatic Control*, 55(10), 2282 – 2297.

Porreca, R., Cinquemani, E., Lygeros, G., and G., F.T. (2010). Identification of genetic network dynamics with unate structure. bioinformatics. *Bioinformatics.*

Roll, J., Bemporad, A., and Ljung, L. (2004). Identification of piecewise affine systems via mixed-integer programming. 40(1), 37–50.

Vidal, R. (2008). Recursive identification of switched ARX systems. *Automatica*, 44(9), 2274 – 2287.

Vidal, R., Chiuso, A., and Sastry, S. (2002). Observability and identifiability of jump linear systems. In *Proc. IEEE Conf. Dec. and Control*, 3614 – 3619.

Weiland, S., Juloski, A.L., and Vet, B. (2006). On the equivalence of switched affine models and switched ARX models. In *45th IEEE Conf. on Decision and Control.*

Westra, R., Hollanders, G., and Tuyls, K. (2007). The identification of dynamic gene-protein networks. In *Springer Lecture Notes in Bioinformatics*, volume 4366, 157–171.